Simulating NC Background from Data

Andrew Culling, Cambridge University HEP Group

UNIVERSITY OF CAMBRIDGE

Motivation

- By putting a CC shower through the Reconstruction chain, this should resemble a NC event.

-By comparing the Data/MC PID distributions for normal and ‘Pseudo-NC’ events. If it accounts for most of the Data/MC discrepancy then it could be used to estimate systematic uncertainty of NC contamination in high (>-0.2) PID region.

- Necessary to first show reasonable agreement between PID variables for NC and CC shws.

Method

- Take NearDet R1_18_2 Selected CC events and select strips from the primary shower and form a new striplist

- It is very important to remove the μ signal (to stop the trkfinder finding the original track) while still including the core of the shw near the vtx. Use the strips from the fitted track as this is more accurate in the shw.

-Strips which are shared between the track and the shower are only used if their charge is > 2.5 MIP and shared strips that are included have their charges reduced by 1.416/dcosz MIP

-The new strip list is then given to the Slicer and the rest of the reconstruction chain is run to the end.

Data/MC

DATA: Dec05 Runs 9303, 9322, 9225, 9331,9334

3.97 E18 PoT

Standard Beam Cuts

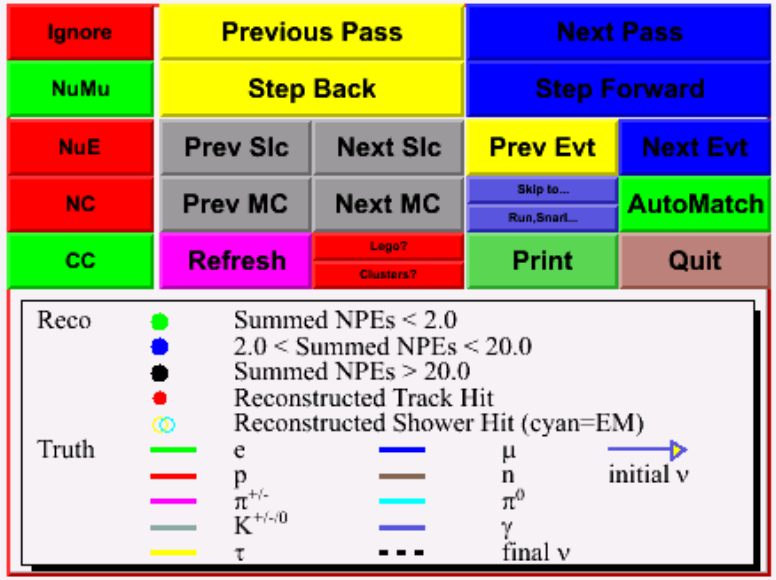
Normal + TrackRemoval

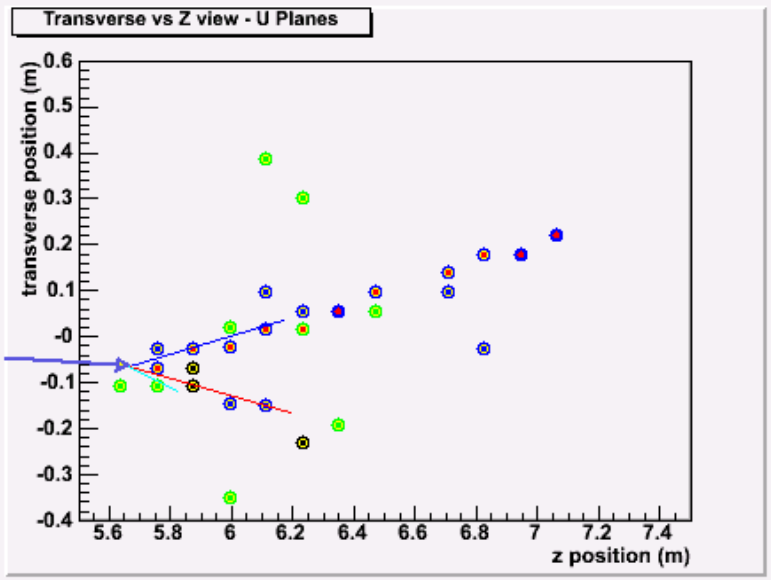
MC: 328 files (3.26 E18) Normal

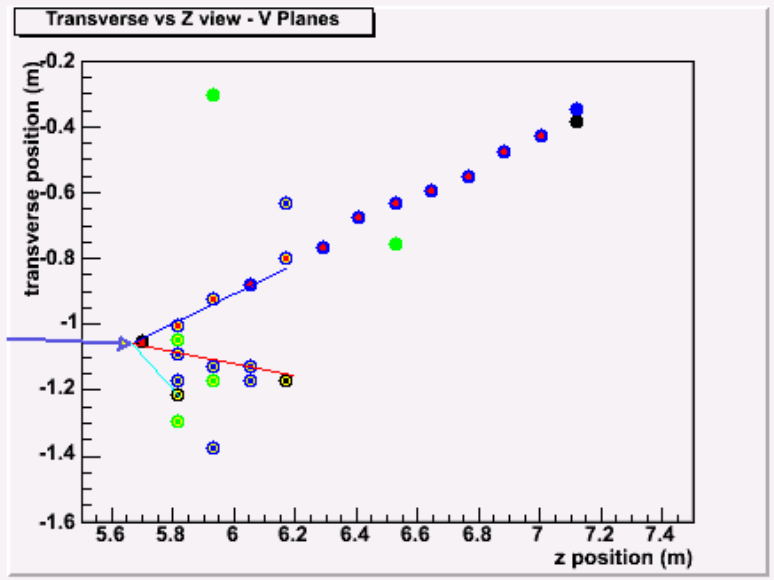
102 files (1.01 E18) TrackRemoval

Initial CC Event

Run: 14301024, Snarl: 119, Slice: 6(/9), Event 4(/7)
Reco - Slice (0.996, 1.000)
#Trks: 1 (0.913, 0.875)
#Shws: 1 (1.000, 0.000)
q/p: 0.310 +/- 0.151, p/q: 3.231
TrkRangeEnergy: 1.013 RecoShwEnergy: 2.502
Vtx: 0.69, -0.80, 5.70
Truth - MC: 8(/26)
Nu ID: 14; NC/CC: 1; Process: 1002
Nu E: 3.110; Mu E*q: -1.692
Mu p: 1.680; Py: 0.68
θ: 0.5040 rad, 28.88 deg
Shw Energy: 1.411657
Vtx: 0.70, -0.79, 5.66

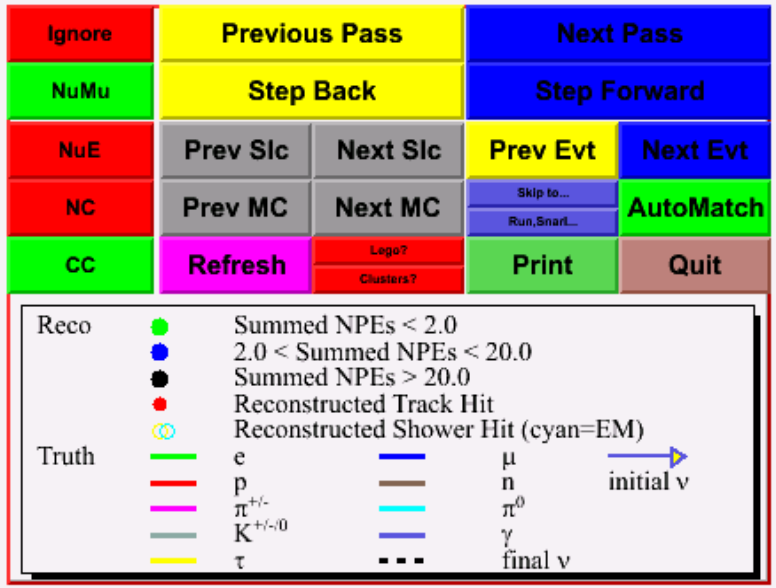


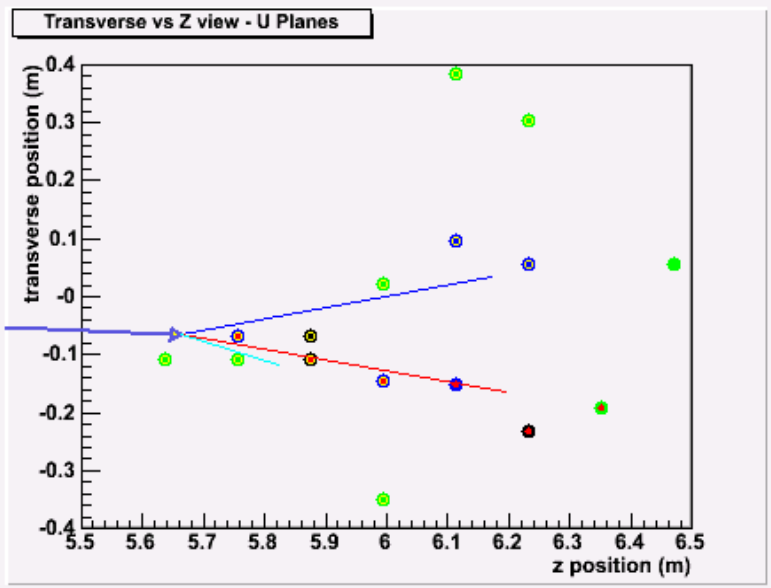


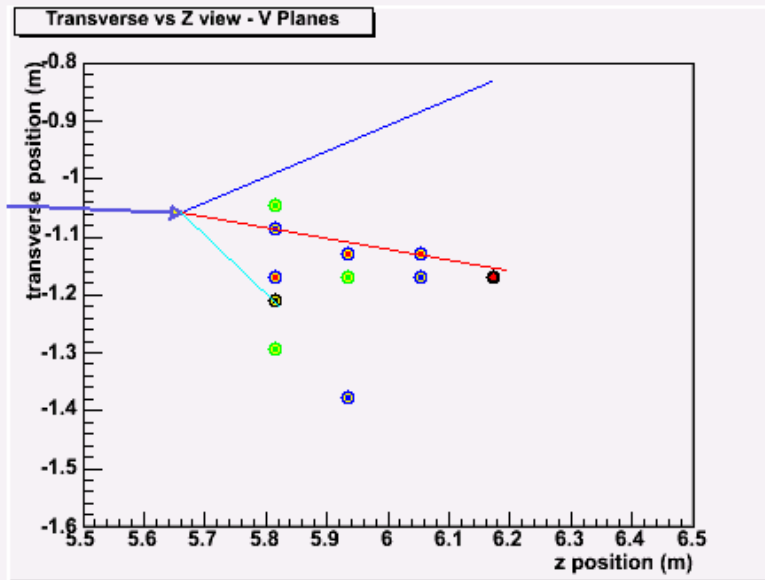


Reprocessed CC Shw

Run: 14301024, Snarl: 119, Slice: 2(/3), Event 2(/3)
Reco - Slice (0.998, 1.000)
#Trks: 1 (0.900, 0.778)
#Shws: 1 (1.000, 0.000)
q/p: 0.881 +/- 0.525, p/q: 1.135
TrkRangeEnergy: 0.468 RecoShwEnergy: 1.219
Vtx: 0.78, -0.88, 5.76
Truth - MC: 8(/26)
Nu ID: 14; NC/CC: 1; Process: 1002
Nu E: 3.110; Mu E*q: -1.692
Mu p: 1.680; Py: 0.68
θ: 0.5040 rad, 28.88 deg
Shw Energy: 1.411657
Vtx: 0.70, -0.79, 5.66







Initial CC Event

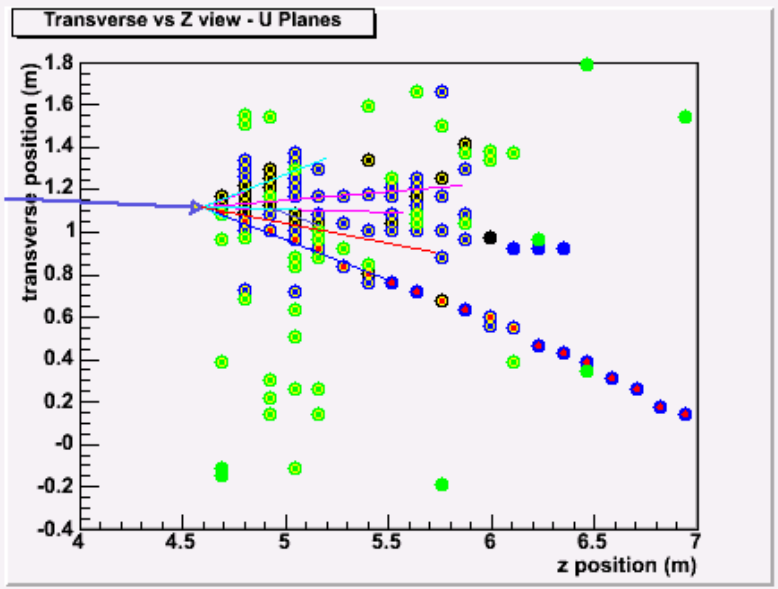
Run: 14301024, Snarl: 164, Slice: 9(/9), Event 6(/7)
Reco - Slice (0.967, 0.997)
#Trks: 1 (0.980, 0.812)
#Shws: 1 (1.000, 0.870)
q/p: -0.256 +/- 0.029, p/q: -3.907
TrkRangeEnergy: 3.102 RecoShwEnergy: 16.093
Vtx: 1.48, 0.09, 4.63
Truth - MC: 7(/33)
Nu ID: 14; NC/CC: 1; Process: 1003
Nu E: 16.904; Mu E*q: -2.895
Mu p: 2.884; Py: -0.92
θ: 0.3236 rad, 18.54 deg
Shw Energy: 14.003059
Vtx: 1.48, 0.09, 4.59

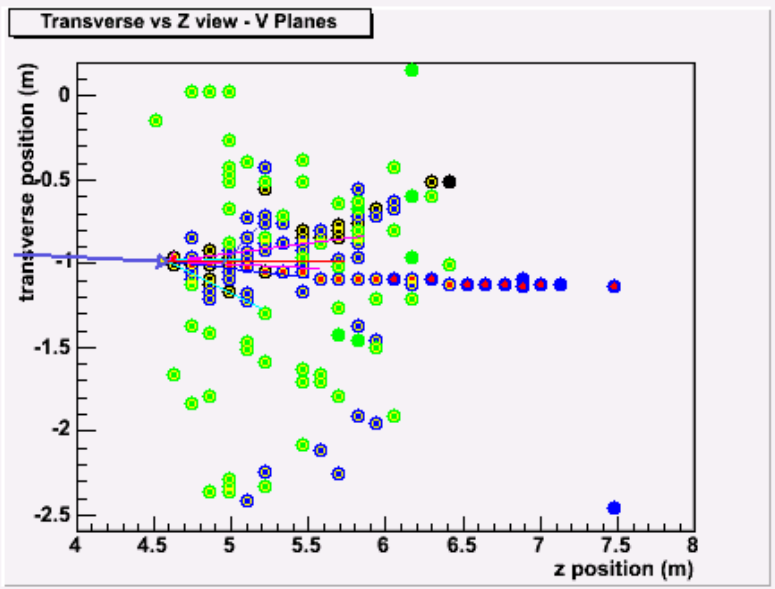
Ignore	Previous Pass		Next Pass	
NuMu	Step Back		Step Forward	
NuE	Prev Slc	Next Slc	Prev Evt	Next Evt
NC	Prev MC	Next MC	Skip to... / Run,Snarl...	AutoMatch
CC	Refresh	Lego? / Clusters?	Print	Quit

Reco
- Summed NPEs < 2.0
- 2.0 < Summed NPEs < 20.0
- Summed NPEs > 20.0
- Reconstructed Track Hit
- Reconstructed Shower Hit (cyan=EM)

Truth
- e
- μ
- p
- n
- $\pi^{+/-}$
- π^0
- $K^{+/-/0}$
- γ
- τ
- final ν

initial ν





Reprocessed CC Shw

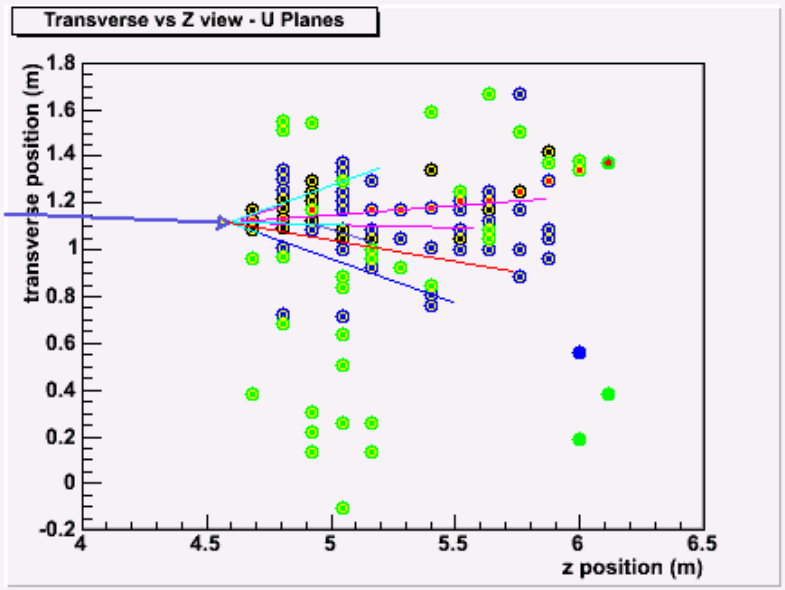
Run: 14301024, Snarl: 164, Slice: 2(/2), Event 2(/2)
Reco - Slice (0.997, 1.000)
#Trks: 1 (0.692, 0.647)
#Shws: 1 (0.997, 1.000)
q/p: 0.615 +/- 0.343, p/q: 1.625
TrkRangeEnergy: 1.153 RecoShwEnergy: 14.448
Vtx: 1.52, 0.08, 4.63
Truth - MC: 7(/33)
Nu ID: 14; NC/CC: 1; Process: 1003
Nu E: 16.904; Mu E*q: -2.895
Mu p: 2.884; Py: -0.92
θ: 0.3236 rad, 18.54 deg
Shw Energy: 14.003059
Vtx: 1.48, 0.09, 4.59

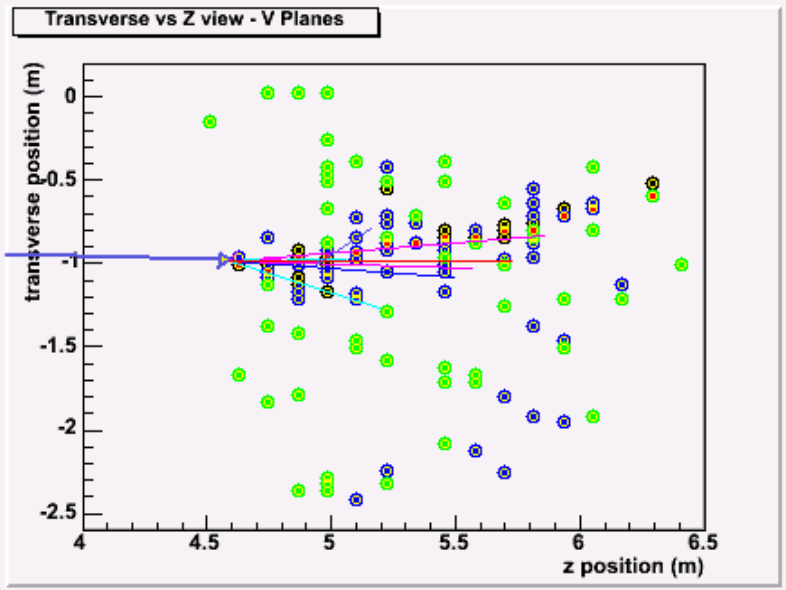
Ignore	Previous Pass		Next Pass	
NuMu	Step Back		Step Forward	
NuE	Prev Slc	Next Slc	Prev Evt	Next Evt
NC	Prev MC	Next MC	Skip to... / Run,Snarl...	AutoMatch
CC	Refresh	Lego? / Clusters?	Print	Quit

Reco
- Summed NPEs < 2.0
- 2.0 < Summed NPEs < 20.0
- Summed NPEs > 20.0
- Reconstructed Track Hit
- Reconstructed Shower Hit (cyan=EM)

Truth
- e
- μ
- p
- n
- $\pi^{+/-}$
- π^0
- $K^{+/-/0}$
- γ
- τ
- final ν

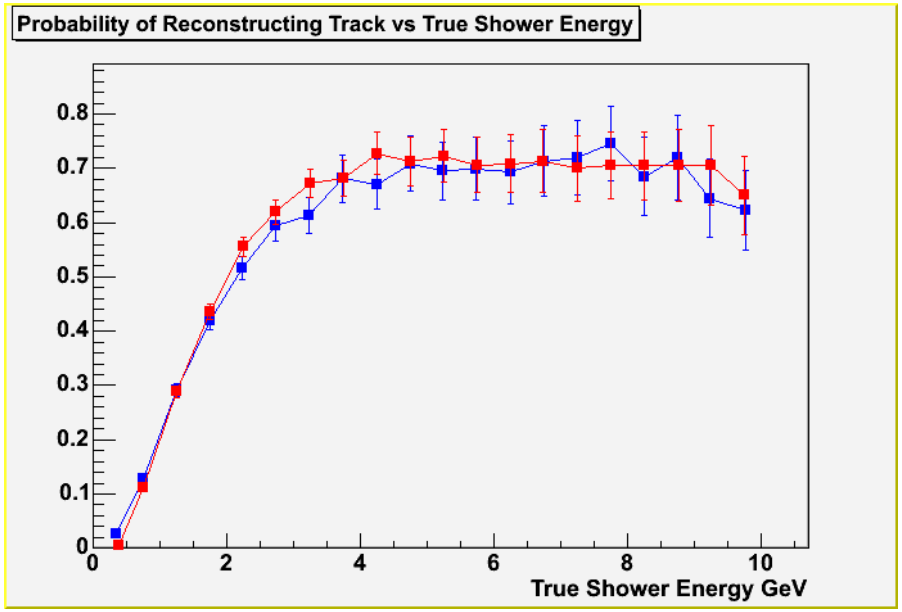
initial ν

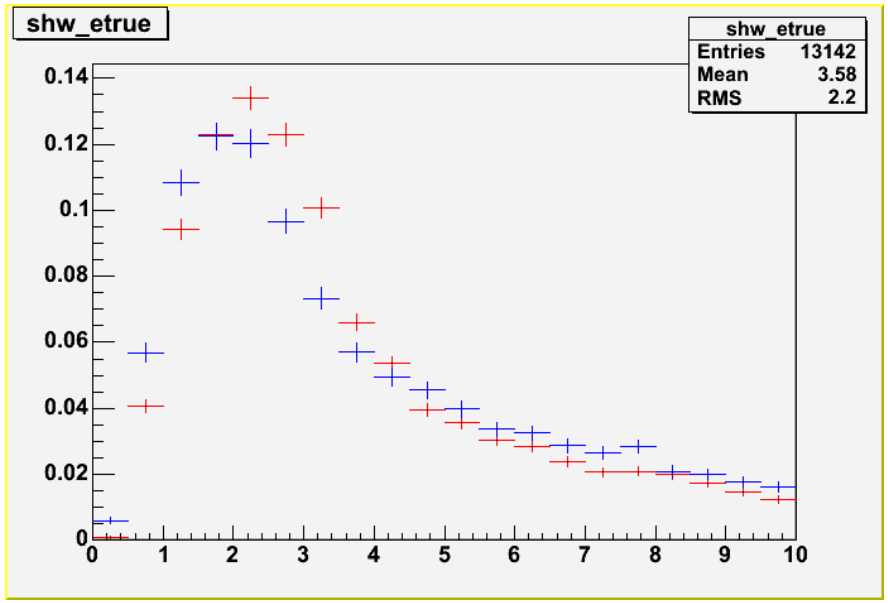




Distributions

Red : NC Blue : CC shws





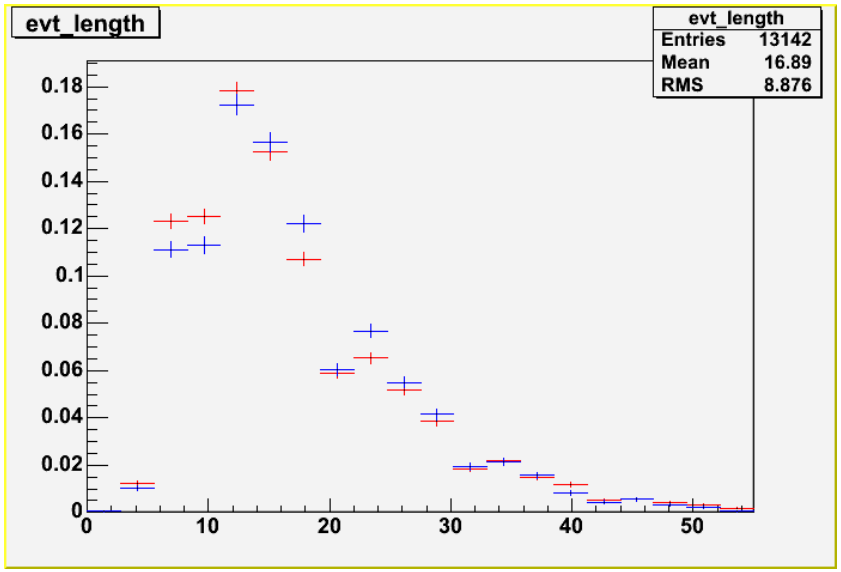
- Reasonable agreement between CC shws and true NC.

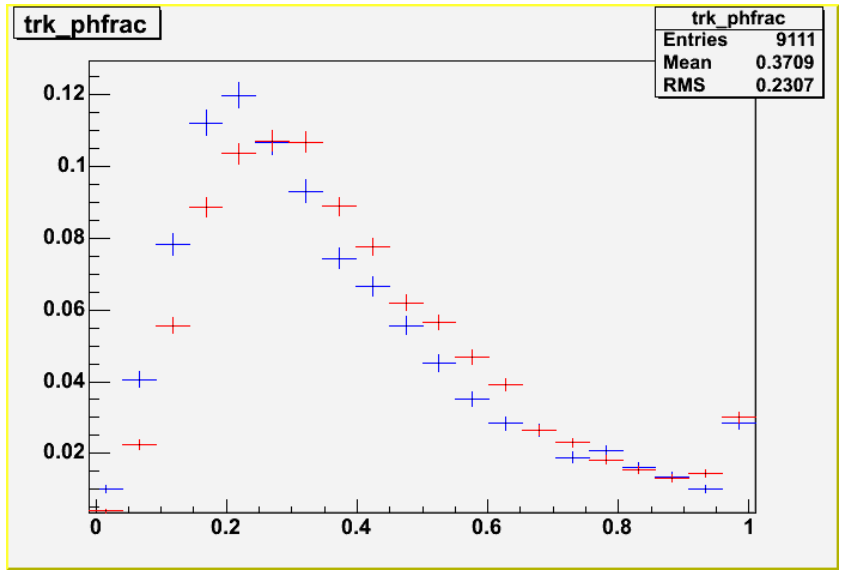
- Initial presence of the track means that many more low energy CC events get through the first pass of TrackSR, so there is a slightly greater fraction of low energy events than with True NC events.

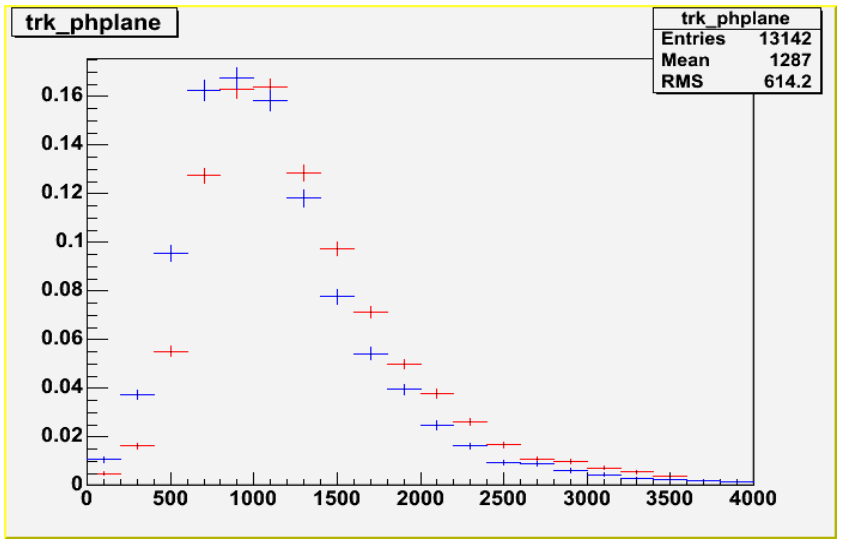
- Have tried reweighting to take into account initial CC/NC shower energy spectra but makes minimal difference to final result.

Distributions

Red : NC Blue : CC shws





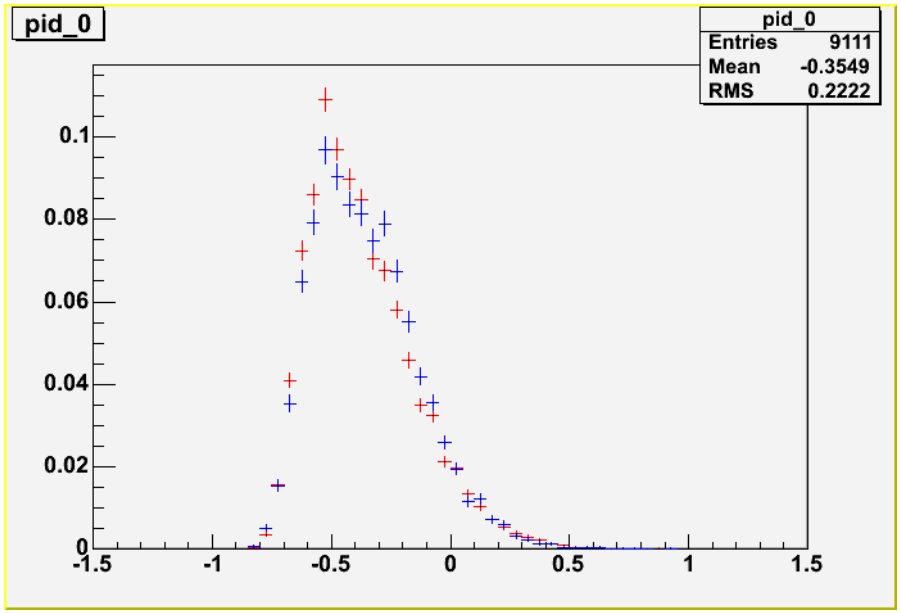


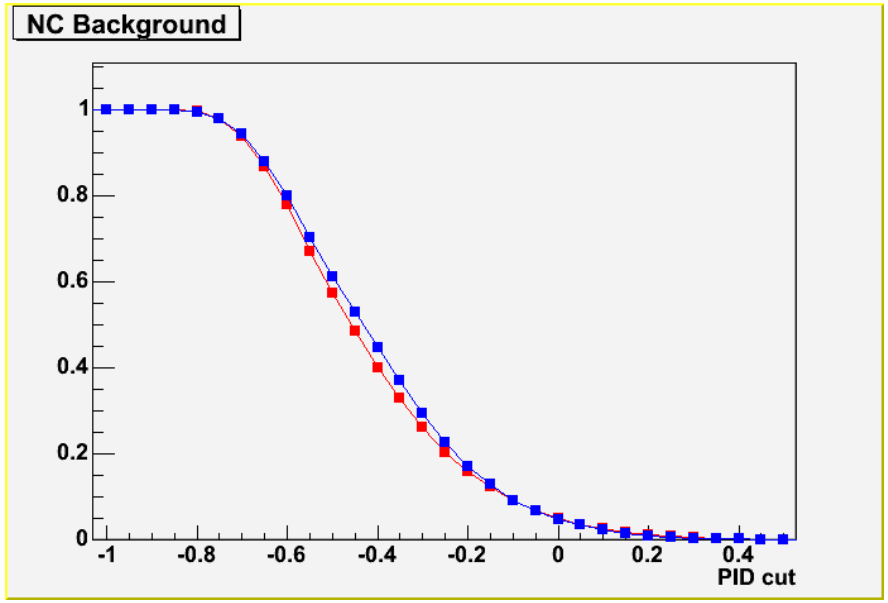
- Reasonable agreement between the PID variables of the two MC samples.

- Differences will be due to incomplete trk charge removal and the excess of low energy shws from the initial CC distribution.

Distributions

Red : NC Blue : CC shws

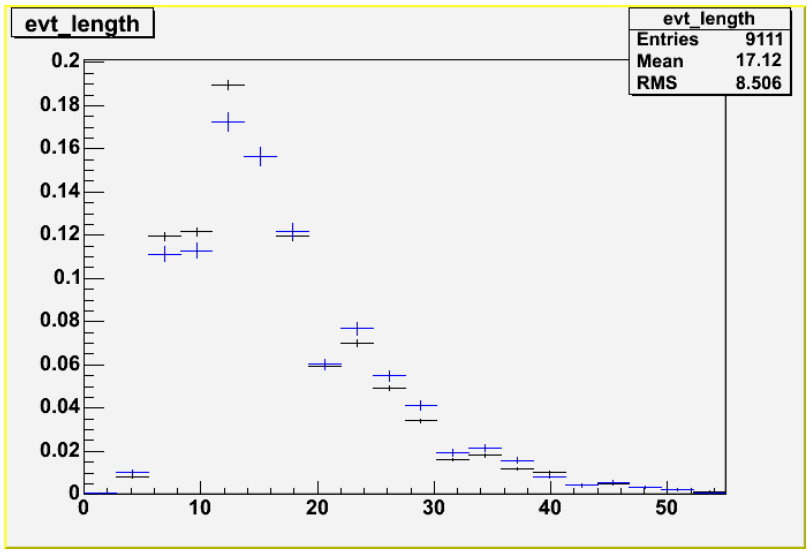


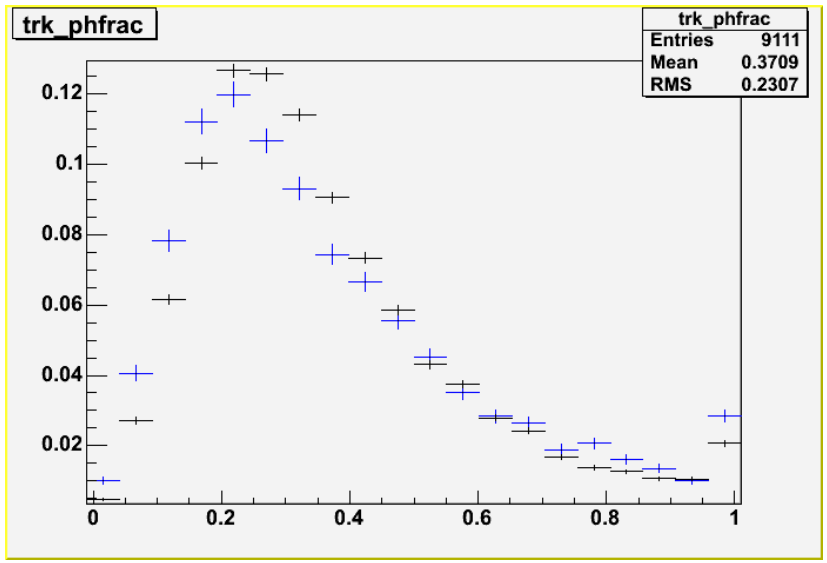


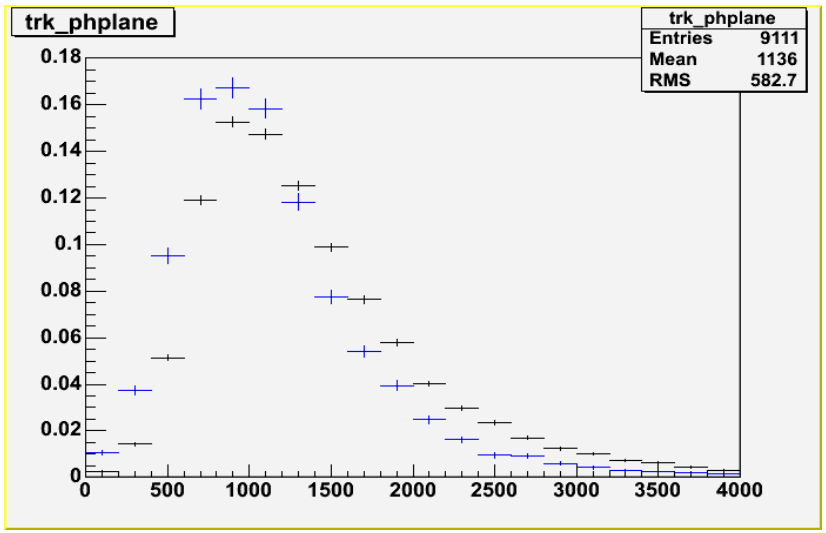
- Reasonable agreement between the PID variables of the two MC samples.

Distributions

Blue : CC shws MC Black : CC shws Data





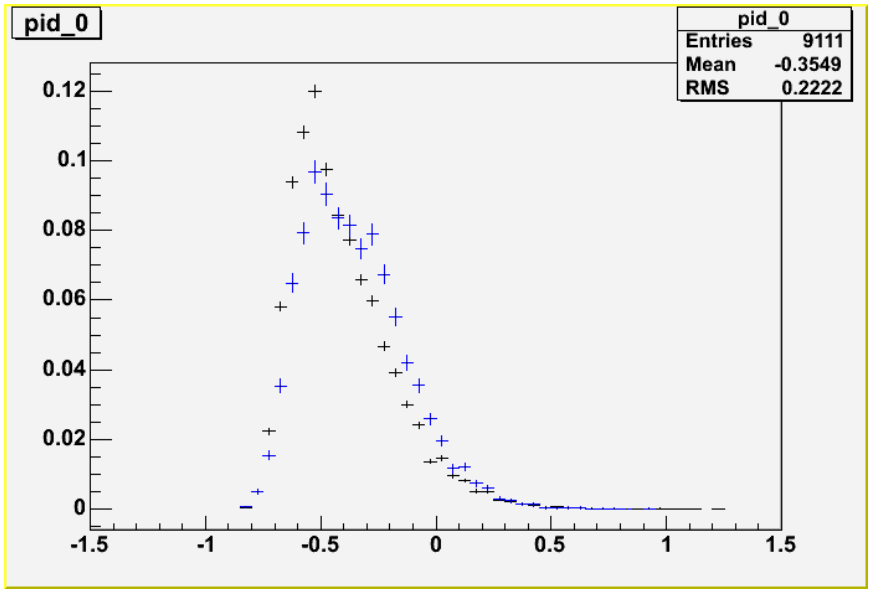


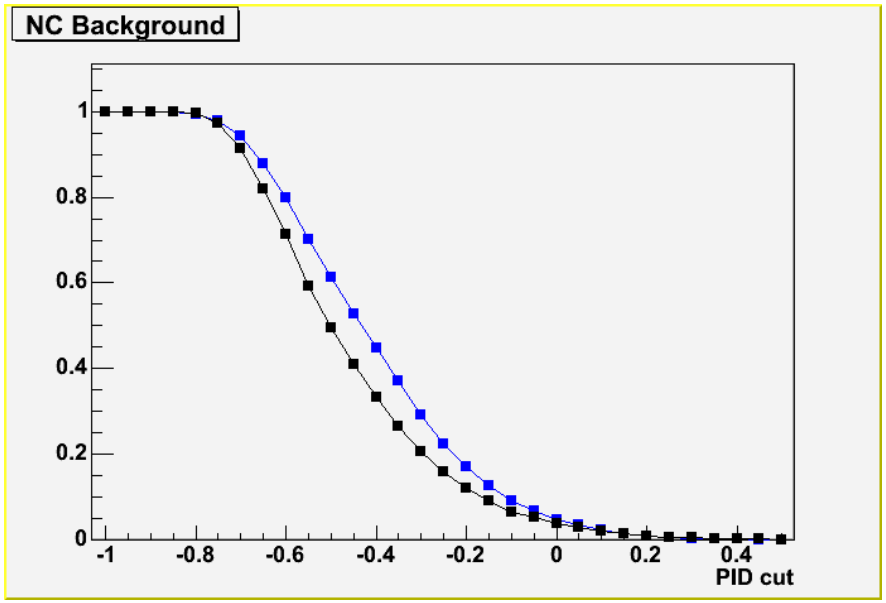
- Resulting PID distributions is now quite different between Data and MC.

- Mainly due to long tail in the trk_phplane dist for data.

Distributions

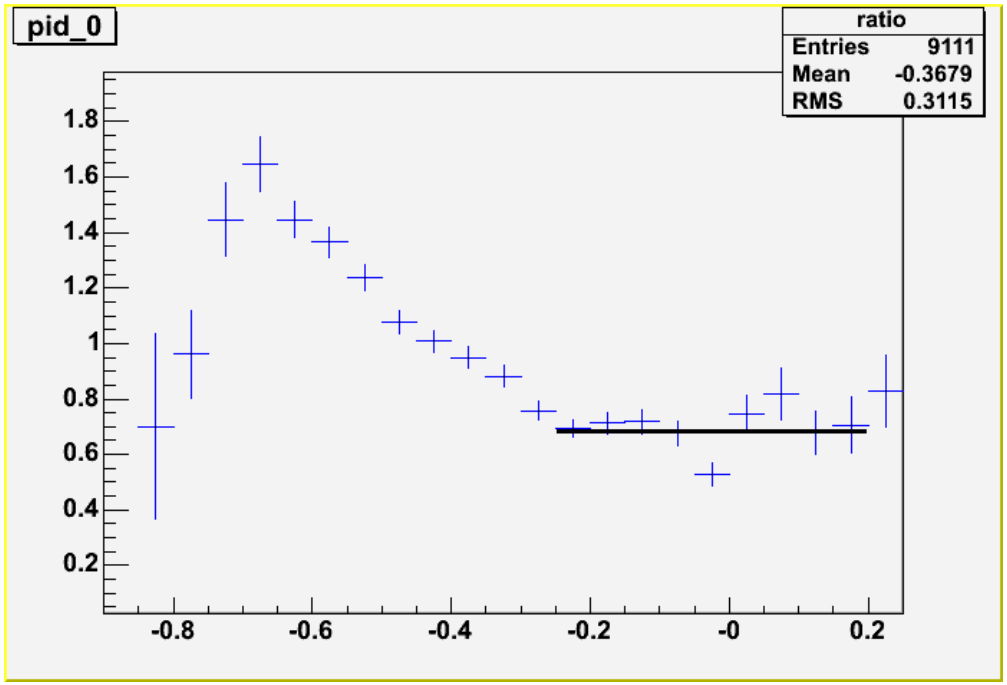
Blue : CC shws MC Black : CC shws Data





- Differences between data and MC CCshws much larger than differences between MC CCshws and NC.

- Plot of ratio of Data/MC vs. PID for CC shws shows there are some substantial variations.

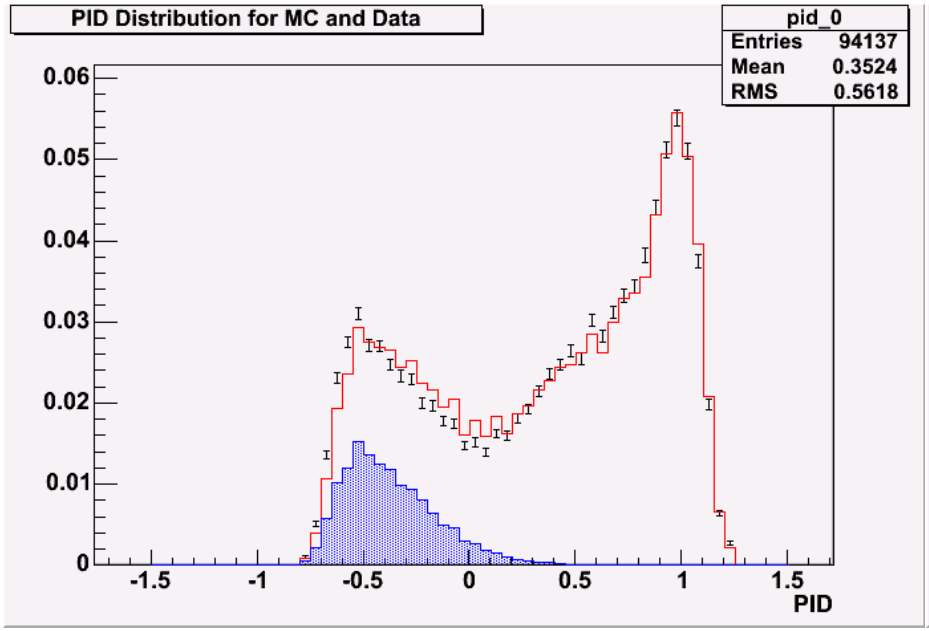


- Seems ~constant ratio (0.68) for PID > -0.25, so can be used to assign a 32% systematic error on knowledge of NC contamination in selected CC sample or could be used as a scaling factor for MC.

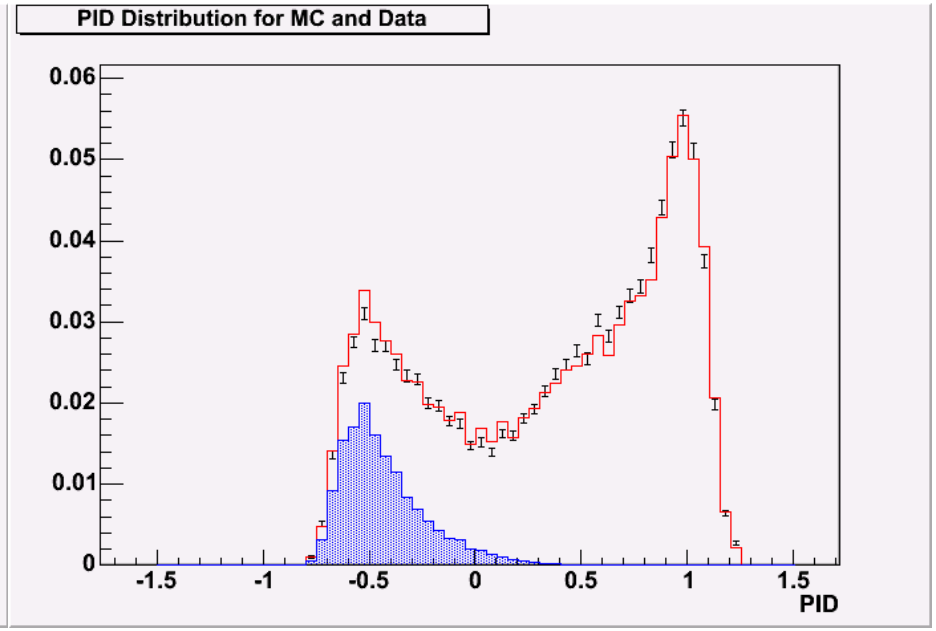
- Make a series of linear fits to describe ratio and apply this to NC component of ND Monte Carlo and compare to data.

Effect of NC Weighting

Before



After



- Appears to be much improved agreement between data/MC. Would suggest that majority of problem with poor data/MC PID agreement is down to poor modelling of hadronic showers.

- χ^2 (for PID<0.2 when scaled up to MC stats) goes from 479/19 to 117/19 when weighting is applied.

Conclusions

- Have shown reasonable agreement between distributions of processed CC shws and true NC events in R1_18_2 NearDet MC.

- Large differences seen for comparison of Data/MC for CC shws.

- Application of observed Data/MC PID ratio to the NC component of ND Monte Carlo greatly improves Data/MC agreement in low PID region. Suggesting that majority of Data/MC difference comes from poor hadronic shower modelling.

- Constant Data/MC ratio for CC shws with PID > -0.25 allows for estimation of systematic uncertainty of NC contamination of CC sample ~ 32% also could be used to scale MC.